

ESERCIZIO

La tabella indagineStudentiFlorida, disponibile sul sito (in formato pdf e xls), riporta i dati di un'indagine condotta su 60 studenti laureati presso la University of Florida, relativamente alle seguenti variabili:

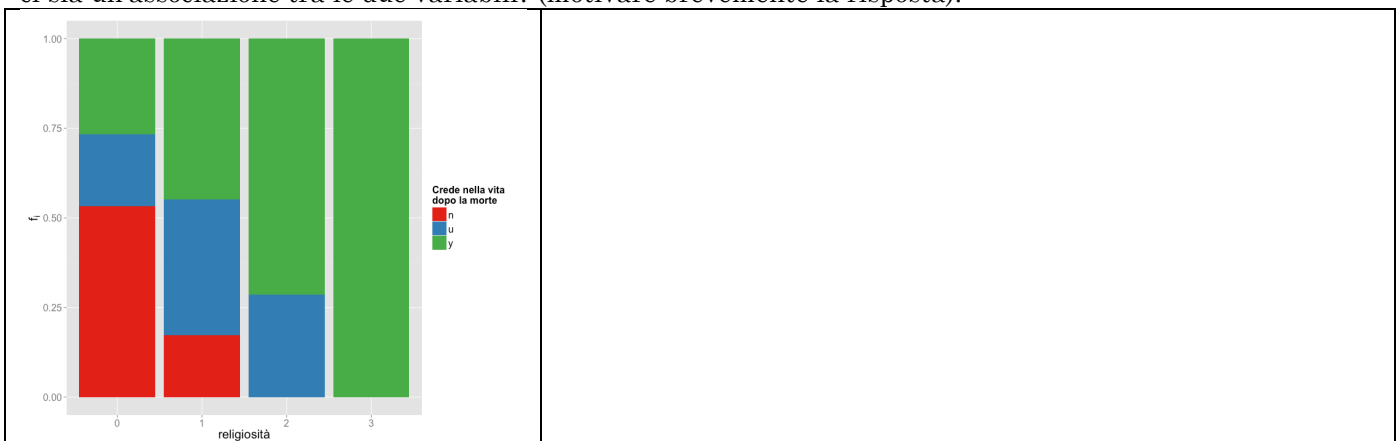
COLONNA	NOME VARIABILE	DESCRIZIONE
1	id	Codice identificativo dello studente
2	genere	Genere dello studente
3	età	Età (in anni compiuti)
4	punteggioHighSchool	Punteggio ottenuto alla scuola superiore (scala da 1 a 4)
5	punteggioCollege	Punteggio ottenuto al College (scala da 1 a 4)
6	distanzaCampus	Distanza (in miglia) del campus dalla città di provenienza dello studente
7	distanzaAula	Distanza (in miglia) della classe dall'attuale luogo di residenza
8	oreTVsettimana	Numero medio di ore settimanali passate a guardare la TV
9	oreSportSettimana	Numero medio di ore settimanali dedicate allo sport
10	quotidianiPerSettimana	Numero di volte per settimana che si legge un quotidiano
11	AH	Numero di persone conosciute morte per AIDS o sieropositive
12	vegetariano	Regime di alimentazione vegetariano (yes, no)
13	affiliazionePolitica	Affiliazione politica (D = Democratico, R = Repubblicano, I = Indipendente)
14	ideologiaPolitica	Ideologia politica (1 = molto liberale, 2 = liberale, 3 = leggermente liberale, 4 = moderato, 5 = abbastanza conservatore, 6 = conservatore, 7 = molto conservatore)
15	religiosità	Partecipazione ad eventi religiosi (0 = mai, 1 = occasionalmente, 2 = la maggior parte delle settimane, 3 = ogni settimana)
16	opinioneAborto	Opinione circa la legalità dell'aborto nei primi tre mesi di gravidanza (yes, no)
17	azioniDiscriminazione	Impegno in azioni contro le discriminazioni (si, no)
18	vitaMorte	Crede nella vita dopo la morte (si -y , incerto - u , no - n)

Fonte: A. Agresti and B. Finlay, *Statistical Methods for the Social Science*, Prentice Hall, 2008

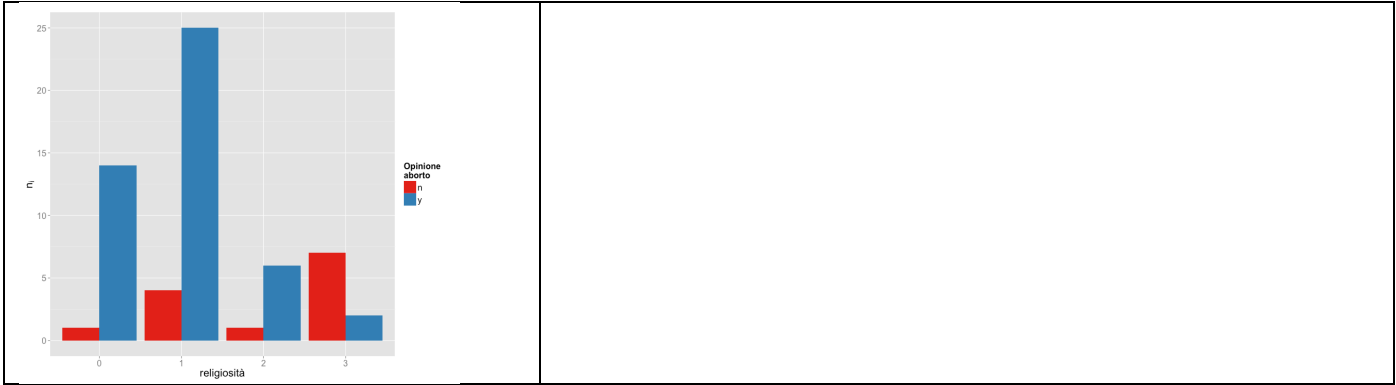
1) Il seguente diagramma a barre riporta la distribuzione della variabile **religiosità**. Commenta brevemente la distribuzione:



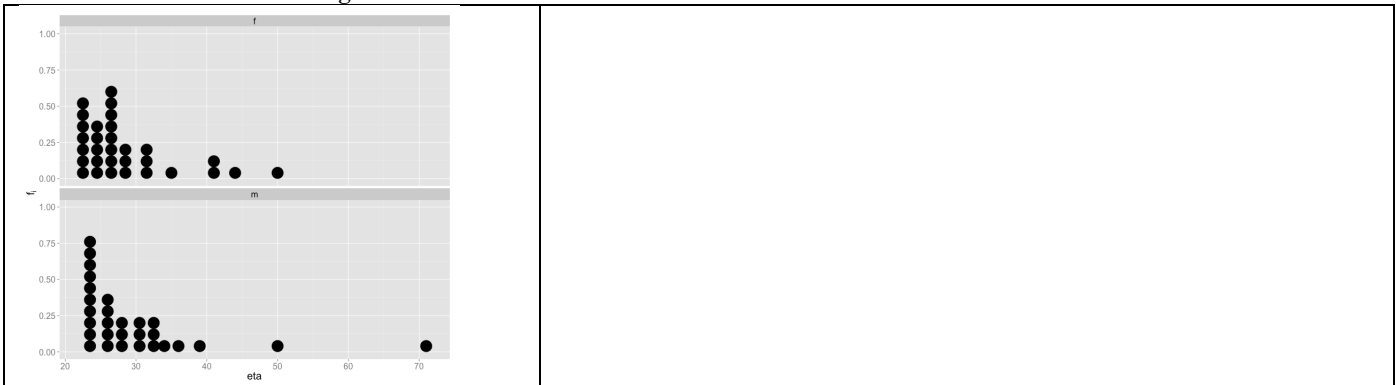
2) Il seguente diagramma a barre impilate riporta la distribuzione della variabile **religiosità** condizionata sulla variabile **crede nella vita dopo la morte**. Dall'analisi del grafico secondo te è possibile ipotizzare che ci sia un'associazione tra le due variabili? (motivare brevemente la risposta):



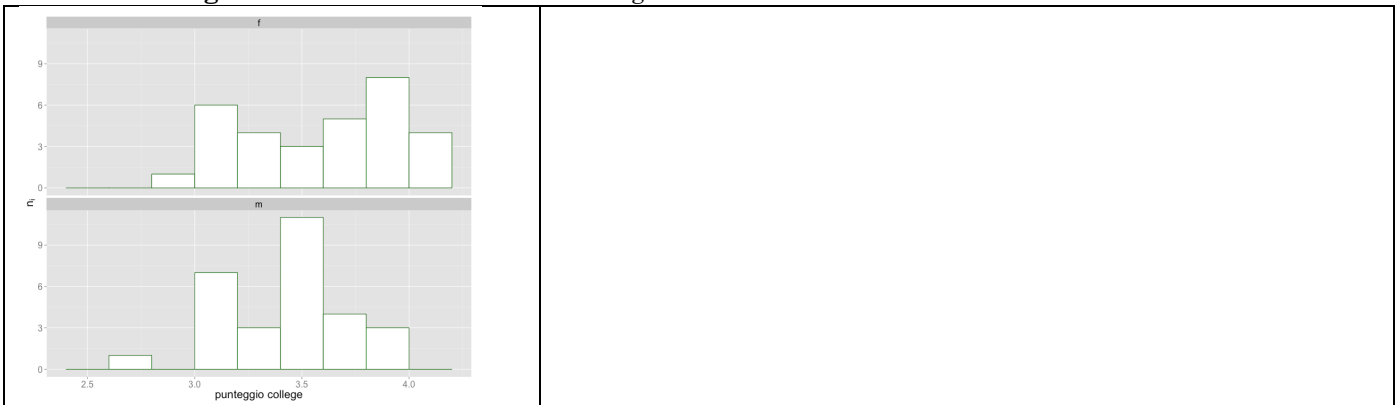
3) Il seguente diagramma a barre accostato riporta la distribuzione della variabile **religiosità** condizionata sulla variabile **opinione aborto**. Dall'analisi del grafico secondo te è possibile ipotizzare che ci sia un'associazione tra le due variabili? (motivare brevemente la risposta):



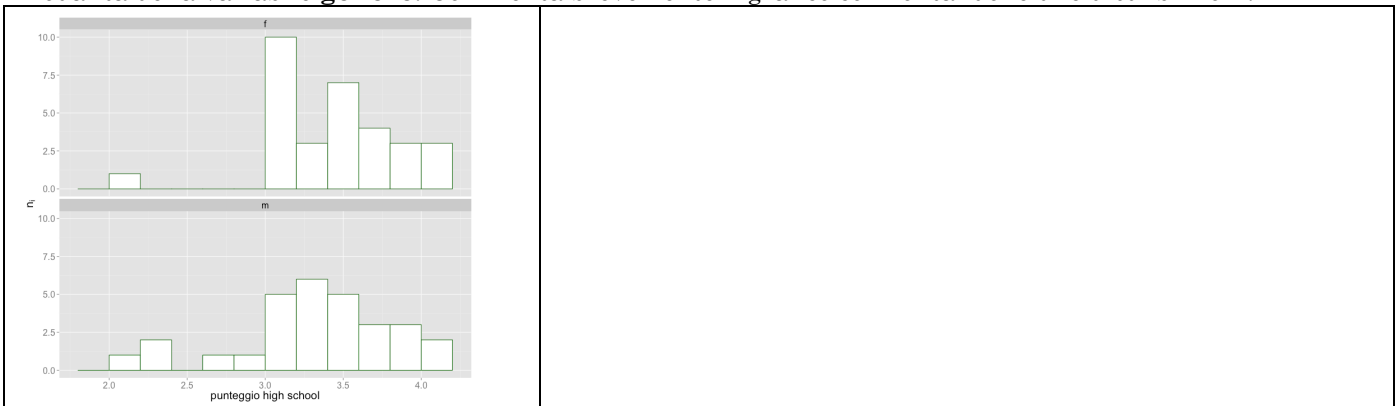
4) I seguenti dotplot riportano la distribuzione della variabile **età** per le due modalità della variabile **genere**. Commenta brevemente il grafico confrontando le due distribuzioni:



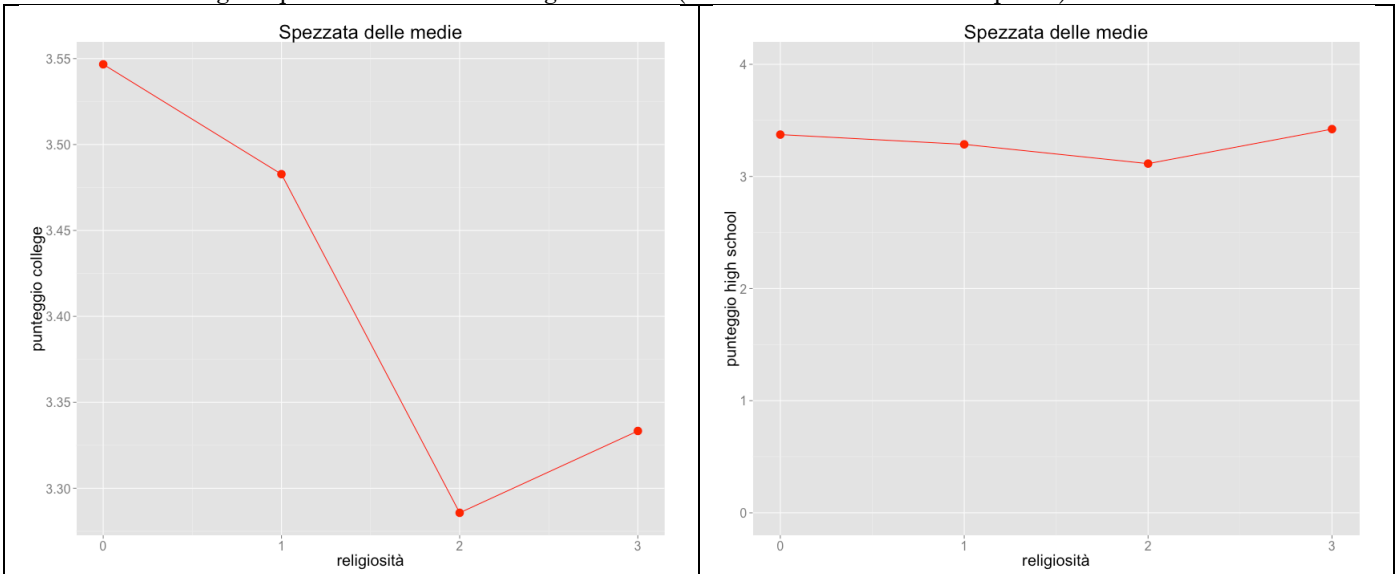
5) I seguenti istogrammi riportano la distribuzione della variabile **punteggio college** per le due modalità della variabile **genere**. Commenta brevemente il grafico confrontando le due distribuzioni:



6) I seguenti istogrammi riportano la distribuzione della variabile **punteggio high school** per le due modalità della variabile **genere**. Commenta brevemente il grafico confrontando le due distribuzioni:



7) I seguenti grafici riportano le spezzate delle medie per le variabili **punteggio college** (grafico sulla sinistra) e **punteggio high school** (grafico sulla destra) condizionati sulle quattro modalità della variabile **religiosità**. Dall'analisi dei due grafici ti sembra che la religiosità di una persona influenzi più il punteggio ottenuto al college o quello ottenuto alla high school? (motiva brevemente la risposta)



Empty box for student response to question 7.

8) Costruisci la tabella di frequenze per classi della variabile **punteggio high school** utilizzando quattro classi di uguale ampiezza:

Empty box for student response to question 8.

9) Rappresentare graficamente la tabella ottenuta al punto precedente utilizzando un istogramma:

Empty box for student response to question 9.

10) Costruisci la tabella di frequenze per classi della variabile **punteggio high school** utilizzando quattro classi di uguale frequenza:

11) Rappresentare graficamente la tabella ottenuta al punto precedente utilizzando un istogramma:

12) Qual è la principale differenza nel procedimento che hai seguito nel costruire i due istogrammi ai punti 9 e 11? (motiva brevemente la risposta):

NOTA BENE: per rispondere ai seguenti esercizi è necessario studiare le dispense relative agli indici di dispersione e di forma disponibili sul sito nella sezione **Materiale Utile**.

13) La seguente tabella riporta la distribuzione della variabile **religiosità**.

religiosità	ni
0	15
1	29
2	7
3	9
TOT.	60

Calcolare i seguenti indici di sintesi:

Moda	
Mediana	
Indice di eterogeneità	
Indice di dispersione	

Indice di forma	
-----------------	--

14) La seguente tabella riporta la distribuzione di frequenza per la variabile **ideologia politica**:

ideologia politica	ni
Molto conservatore	2
Conservatore	4
Abbastanza conservatore	4
Moderato	10
Leggermente liberale	6
Liberale	24
Molto liberale	8
TOT.	60

Trattando la variabile come una variabile ordinale (dove Molto conservatore = 7 è stato gestito come il livello più basso della scala e Molto liberale = 1 è stato gestito come il livello più alto della scala) è possibile ottenere i seguenti valori per gli indici di variabilità e forma:

Indice di eterogeneità	Indice di dispersione	Indice di dispersione prime 30 osservazioni	Indice di dispersione ultime 30 osservazioni
0.769	1.758	1.440	0.391

Considerando tali risultati e quelli ottenuti al punto precedente, confronta le due variabili religiosità ed ideologia politica in termini di eterogeneità, dispersione e forma (motivando brevemente le risposte):

Eterogeneità	
Dispersione	
Forma	